THE UNIVERSITY of York

Discussion Papers in Economics

No. 2009/14

Does Education Reduce Blood Pressure? Estimating the Biomarker Effect of Compulsory Schooling in England

By

Nattavudh Powdthavee

Department of Economics and Related Studies
University of York
Heslington
York, YO10 5DD

Does Education Reduce Blood Pressure? Estimating the Biomarker Effect of Compulsory Schooling in England

Nattavudh Powdthavee

University of York

July 2009

Abstract

This paper is the first of its kind to estimate the exogenous effect of schooling on reduced blood pressure and the incidence of hypertension. Using the changes of the minimum school-leaving age in the United Kingdom from age 14 to 15 in 1947, and from age 15 to 16 in 1973, as instruments, the IV-probit estimates imply that completing an extra year of schooling reduces the probability of developing subsequent hypertension by approximately 5%-11% points. The correct estimates of the LATE for schooling indicate the presence of a large and negative bias in the least square/probit estimates of schooling-health relationship.

Keywords: blood pressure; compulsory schooling; biomarker; IV; hypertension; health

JEL: H1, I1, I2

Ever since Michael Grossman's seminal work on the demand for health in the early 1970s (Grossman, 1972), researchers have routinely tried to estimate the effects of different socio-economic variables on a variety of health outcomes. So far, years of schooling has stood out as one of the largest determinants of health. This is true whether health levels are measured by self-assessed health status, mortality rates, risky health behaviors, morbidity, physiological measures, and mental well-being (Grossman, 1975; Berger and Leigh, 1989; Kenkel, 1991; Deaton and Paxson, 2001; Oreopoulos, 2007; Blanchflower and Oswald, 2008).

There are, however, two major practical problems associated with the estimation of the schooling effect on health. The first corresponds to the classic estimation bias associated with ordinary least squares (OLS) estimates of the return to schooling. In the economics literature, researchers frequently use the instrumental variables (IV) approach to address the endogeneity of schooling decisions. A valid instrumental variable, which determines whether an individual receives more years of schooling, but does not determine other factors that affect the outcome of interest, can overcome estimation biases that often arise when using the OLS method. Yet, according to Guido Imbens and Joshua D. Angrist (1994) and Philip Oreopoulos (2006), many of the instruments used in the return to schooling literature – e.g., distance from home to college (Card, 1995), restrictive compulsory schooling law (Angrist and Krueger, 1991), and regional spending on education in regions where the individual was still a student (Berger and Leigh, 1989) – only affect a small fraction of the general population. As a result, many of the IV estimates produced in the literature are only approximations of the average treatment effects among a small group of people who happened to be exposed to the instruments (Card, 2001).

The second practical problem concerns the existing measures of health outcomes frequently used in the studies conducted by economists concerning the determinants of health. While there are valid reasons for using self-assessed health, mortality rates, morbidity, and risky health behaviors such as smoking and drinking as proxies for health, these variables can only, at best, be considered by the medical professions as indirect indicators of someone's underlying health. They do not, for example, possess the same clinical properties as such biomarkers as blood pressure, cortisol levels, cholesterol levels, or heart rate, which are normally used by clinicians to measure someone's biologic state. For example, self-reported health, which is the most commonly used measure of health outcomes in economics, is subject to a variety of potential measurement biases and interpersonal comparability problems (for a recent review, see Powdthavee, 2009). The same holds true for other self-reported health problems. Mortality rates, morbidity, and disability index may fair better as measures of health, given that each has the required quality of being ‘objective’. Yet both are indicators of the extreme cases of ill-health and physical disablement, which will reflect only a small fraction of the nation’s population. Because nationally representative surveys that contain both biomarker readings and years of schooling are scarce, econometric evidence on the biomarker effect of schooling is virtually non-existent.

This paper aims to fill that research void by using the unique Health Survey for England data set, which combines both interviews and physical examinations, to study the effects of schooling on adult blood pressure readings in England. Following Harmon and Walker (1995) and Oreopoulos (2006, 2007), I will rely on exogenous changes in the amount of schooling received by individuals caused by the raising of the minimum school-leaving age in the United Kingdom (which has occurred twice over the age-spread of those over the age of 16 in the English data set) to provide instruments for schooling. Because the changes of minimum school-leaving age affected virtually everyone who would have left school at an earlier age before the introduction of the law, the IV estimates of the schooling effects on blood pressure is likely to come close to the average treatment effects that apply to the whole population (see Oreopoulos, 2006). In addition to the IV approach, I also adopt a regression discontinuity (RD) approach in order to illustrate the average educational attainment and adult blood pressure just before and after the introduction of the minimum school-leaving age law.

There are empirically good reasons to use blood pressure readings as the main biomarker for the nation’s underlying health and well-being. Frequently recorded as a continuous variable, blood pressure has been shown in the medical literature as the single most important predictor of heart disease and death from heart disease (Hofman, Feinleib, Garrison, and van Laar, 1981; Fraser, 1986; Wilson and colleagues, 1998), which is also the current biggest killer in America and the UK. Blood pressure has also been used as a biomarker for stress and general psychological well-being in the work by David Blanchflower and Andrew J. Oswald (2008, 2009).

Section I outlines the analytical framework of Grossman’s (2000, 2005) human capital model and its implication on the demand for health. The model provides a useful backdrop to the relationship between schooling and health. Second II gives a brief account of previous empirical evidence on the effects of schooling on health. Data and empirical strategy are discussed in Section III. Section IV presents both RD and IV results. Section V concludes.

I. Analytical framework

A. *Basic human capital model*

Following Grossman's (2000, 2005) theories on human capital and the demand for health, individuals have two motives for demanding health: consumption and investment. As a consumption commodity, health is a direct source of utility. As an investment commodity, it determines the total amount of time in a period that can be allocated to work in the market and to the production of commodities in the nonmarket sector. Let the intertemporal utility function of an average individual be

$$U = U(\phi H_t, Z_t),\ t = 0,1,...,n, \tag{1}$$

where H_t is the stock of health at age *t* or in time period *t*, ϕ_t is the service flow per unit stock of health, and Z_t is consumption of another commodity. The stock of initial period of health (H_0) is given, but the stock of health at any other age is endogenous. The length of life as of the planning date (*n*) is also endogenous, and death takes place when the stock of health falls below the minimum level required to survive ($H_t \leq H_{\min}$). Therefore, length of life is determined by the quantities of health capital that maximize utility subject to production and resource constraints.

We can think of net investment in the stock of health as equal to gross investment minus depreciation:

$$H_{t+1} - H_t = I_t - \delta_t H_t, \tag{2}$$

where I_t is gross investment and δ_t is the rate of depreciation during the t^{th} period ($0 < \delta_t < 1$). The rates of depreciation are exogenous but dependent on age. Individuals produce gross investment in health and the other commodities in the utility function according to a set of household function of:

$$I_t = I_t(M_t, TH_t; E) \tag{3}$$

$$Z_t = Z_t(X_t, T_t; E), \tag{4}$$

Here, M_t is a vector of inputs or goods purchased in the market that contribute to gross investment in health, X_t is a similar vector of goods that contribute to production Z_t, TH_t and T_t are time inputs, and *E* is the individual's stock of knowledge or human capital exclusive of heath capital. In the most simplistic form, *E* is assumed in Grossman's model to be exogenous or predetermined. The efficiency of production process in the nonmarket or household sector increases with the stock of knowledge, *E* – just as an increase in technology raises the efficiency of the production process in the market sector (Michael, 1972, 1973; Michael and Becker, 1973).

B. *Schooling, productive efficiency and health*

Focusing on how *E* determines the demand for *H*, I follow, for simplistic reason, Grossman's static version of a pure investment model in which health does not enter the utility function directly (Grossman, 2005). This is primarily because, unlike the pure consumption model where health enters individual's utility directly, the pure investment model generates powerful prediction on the implications of *E* on *H* from simple analyses.

In the period of interest, say a year, the total amount of time allocated to market and nonmarket production (*h*) is not fixed. Rather, it is a positive function of health (*H*) because an increase in health reduces the time lost from these activities due to illness or injury ($\partial h / \partial H \equiv G > 0$). Because the output of health has a finite upper limit (8,760 days or 365 days$\times$24 hours per day if the relevant period is a year), the marginal product of health falls as *H* rises ($\partial^2 h / \partial^2 H \equiv G_h < 0$). Health is produced with inputs in the market that contribute to gross investment in health, i.e. medical care, *M*, and the individual's time, *TH*:

$$H = e^{\rho_H S} F(M, TH) \qquad (5)$$

where *F* is linear homogenous in *M* and *TH*, *S* is a measure of the efficiency in the production process, and ρ is a positive parameter. The efficiency variable *S* coincides with the individual's stock of knowledge, *E*, which depends on such additional factors such as the quality of schooling and on the job training, although the focus of this paper will be on the number of years completed formal schooling (see Michael, 1972, 1973). An increase in the number of years completed formal schooling is assumed to raise the marginal product of *M* and *T* by the same percentage (ρ_H).

The individual maximizes $Wh - \pi_h H$, where *W* is the wage rate and π_h is the marginal or average cost of producing health. The first order condition for optimum *H* is

$$WG = \pi_h. \qquad (6)$$

Using this equation, we can obtain formulas for the optimal percentage chances in the quantities of *H* and *M* caused by a one unit increase in schooling (*S*):

$$\tilde{H} = \varepsilon_H \rho_H \qquad (7)$$

$$\tilde{M} = (\varepsilon_H - 1)\rho_{H,} \quad (8)$$

where

$$\varepsilon_H \equiv -\frac{G}{HG_H},$$

and

$$G_H = \partial G / \partial H .$$

The effects summarized by (7) and (8) hold the wage rate and the price of medical care constant. Note that the marginal product of health care in the production of healthy time is GH_M, where $H_M = \partial H / \partial M$. An increase in schooling raises H_M. With M constant, however, an increase in schooling reduces GH_M if $\varepsilon_H < 1$.

The parameter ε_H is therefore the inverse of the absolute value of the elasticity of the marginal product of health (*G*) with respect to *H*. Because the output of health has a finite limit, ε_H is likely to be smaller than one (Grossman, 2000). Given that this condition holds, an increase in schooling is predicted to lower the quantity of health care demanded but increase the quantity of health demanded.

C. *Allocative efficiency*

An alternative model that explains the link between education and health outcomes is the model of allocative efficiency. In this model, individuals with more schooling are assumed to have more information about the true nature of their production function (see, e.g., Rosenzweig, 1995; de Walque, 2007). For example, more educated individuals may have more knowledge about the harmful effects of smoking, drinking, and bad diet on their health stock. In addition, they may also respond to new knowledge more rapidly, which leads to an observable correlation between schooling and health outcome variables. However, one implication of the allocative efficiency model implicitly is that the schooling effect in a health production function would be zero once all health-related inputs are accounted for.

In short, whether one adopts the productive efficiency or the allocative efficiency model, schooling is predicted to increase the demand for health for the individual.

II. Previous evidence on schooling and health outcomes

Much of the previous attempts to estimate the effect of schooling on health outcomes have used measures of self-reported health as proxies for individual’s health stock. Using a nationally representative American data set, Grossman (1972) is one of the first to document positive correlations between year of schooling and subjective health. Wagstaff (1986) employs the 1976 Danish data to show that schooling is positively correlated with a measure of good health indicated by having low combined scores of such self-reported health problems as physical mobility, respiratory problems and presence of pain. Erbsland, Reid, and Ulrich (1995) find similar relationships between schooling and self-reported health problems in a nationally representative German data set. Using a 1993 Dutch data set of men and women who were sixth grade pupils in 1953 in the province of Noord-Brabant, Hartog and Oosterbeek (1998) show that schooling have a positive relationship with self-rated health, even after controlling for IQ and parents' schooling among other variables. Gerdtham and Johannesson (1999) obtain similar findings in their subjective health equation using the 1991 Swedish micro data. More recently, Case, Fertig, and Paxson (2005) use the 1958 British National Child Development Study to show that self-reported health of males at age 42 is significantly correlated with the number of years completed formal schooling. The schooling coefficient is positive and statistically significant even in models that include self-rated health at age 23 and 33. For a more extensive review on the relationship between schooling and subjective health, see Grossman (2005).

There have also been attempts by economists to link the effects of schooling on other more objective health outcomes such as obesity and mortality. For example, Chou, Grossman, and Saffer (2004) find using the American Behavioral Risk Factor Surveillance System that schooling has a negative and statistically well-defined relationship with adult obesity. With respect to mortality, Grossman (1975) shows schooling to have a positive and statistically significant effect on the probability of survival for the middle-age white males in the NBER-Thorndike sample. More recently, Deaton and Paxson (2001) conclude using two American data sets that schooling has a negative effect on mortality for persons under the age of 60 as well as for person above that age.

Consistent with the prediction made by the allocative efficiency theory, there is also some evidence of the beneficial effects of schooling on health behaviors. Using data from the Health Promotion/Disease Prevention Supplement to the 1995 National Health Interview Survey, Kenkel (1991) demonstrates schooling to have a negative and significant relationship with smoking and heavy drinking. Goldman and Smith (2002) find that more educated HIV/AIDS patients are more likely to adhere to therapy than their less educated counterpart. Similarly, de Walque (2007) finds that an increase in the level of exposition about the dangers of the HIV/AIDS epidemic supplied by the prevention programs in Uganda in 1990 has resulted in a significant drop in the risk of being HIV positive among young individuals in 2000.

Schooling is unlikely, however, to be exogenous. There are a variety of sources of bias associated with estimation of the schooling effect on health. First, causality may also run in reverse from health to schooling, i.e. healthier students may be more efficient producers of additional human capital via more years of formal schooling, which implies that estimates of the schooling effect on health will be biased upward. Second, there may be omitted third variables such as ability (Angrist and Kruger, 2001; Card, 2001), heritable endowments (Behrman and Rosenzweig, 2002), and time-preference (Fuch, 1982) that influence both schooling and health outcomes. One could imagine, for example, that people who are more future oriented (i.e. those who desire more leisure at older ages) will stay in school for longer, work more at younger ages, as well as have higher levels of health during most stages of the life cycle. Thus, the effect of schooling will be biased upward if one fails to control for time-preference. A third source of potential bias is measurement error, which can bias the estimated schooling effect toward zero (Blackburn and Neumark, 1995).

Previous work on the estimation of the schooling effect on health has mainly dealt with the endogeneity issue using either the IV method or the quasi-experimental approach. Using the state of residence in childhood as instruments for schooling, Leigh and Dhir (1997) find schooling to be negatively related to an index of disability. However, there is little difference in size of the schooling coefficient whether it is treated as exogenous or endogenous. The same IV method is used by Sander (1995) to estimate the causal effect of schooling on smoking in the 1986-1991 National Opinion Research Center’s General Social Survey. Applying parents’ education, rural residence at age 16, region of residence at age 16, number of siblings as instruments for schooling, Sander finds the schooling effects on the probability of quitting smoking estimated by probit and IV-probit estimators to be virtually the same.

More recently, Lleras-Muney (2005) uses compulsory education laws from 1915 to 1939 as instruments for education in the adult mortality equations. When treating schooling as exogenous, she finds the IV estimates on the schooling effect on adult mortality to be negative and significantly larger than the ones obtained by OLS. Adams (2002) demonstrates using the same instruments as Lleras-Muney (2005) that the schooling effect on self-assessed health is much larger in the IV equations than in the OLS equations. Similar results are also obtained by Arendt (2005) when two compulsory schooling reforms in Denmark are used to address the endogeneity of schooling in self-assessed health equations. Using changes in compulsory schooling in Sweden which implemented randomly and in stages by municipalities in the 1950s to instrument for schooling, Spasojevic (2003) reports positive schooling effects on a constructed index of bad health and an index of body mass index (BMI) in the healthy range, although the effects are only significant when using one-tailed tests. In short, most studies do not find significant differences in the estimated schooling effects on health whether the schooling variable is treated as exogenous or endogenous.

One important issue concerning the use of IV method to estimate the return to investment in human capital is that the only effect we can be sure that this method estimates is the local average treatment effect (LATE), i.e. the average treatment effect (ATE) among those who alter their status because they react to the instrument (Imbens and Angrist, 1994). In many cases, the instruments used in the return to schooling literature only apply to a small fraction of the population (see, e.g., Angrist and Kruger, 1991). What this implies is that many of the IV estimates will only approximate the ATE among a small and peculiar group rather than the general population, whereas OLS estimates, in the absence of omitted third variables and measurement error problems, approximate ATE among everyone (Card, 2001). For example, as in the aforementioned study by Spasojevic (2003), because school reforms in Sweden took place randomly and in stages by municipalities, it is possible that her IV estimates only approximate the average treatment effects among students who happened to be residing in these municipalities when the changes took place.

Perhaps one of the more successful instruments used to estimate the market returns to education in recent times which, when implemented, will produce the IV estimate that is closest to the ATE for the general population as possible, is the changes in compulsory schooling law in the UK (see Harmon and Walker, 1995; Oreopoulos, 2006). This is simply because the introduction of the minimum school-leaving age in the UK – from age 14 to 15 in 1947, and from age 15 to 16 in 1973 – affect virtually *everyone* who would have left school at age 14 prior to 1947, and similarly for those who would have left school at age 15 prior to 1973. The dramatic effect of the introduction of such laws on the amount of schooling received by the general population means that the estimated local treatment effects of education will come close to mirroring population average treatment effects (Oreopoulos, 2006).

With respect to the estimation of the schooling effect on health, Philip Oreopoulos (2006, 2007) is among the first to use nation-wide minimum school-leaving age law in the UK to estimate the LATE for schooling on self-assessed health in the nationally representative General Household Survey. Although he finds consistent evidence that education improves self-assessed health (a one-year increase in schooling raises the probability of individuals reporting being in good health by 6% points), there is little evidence that the IV estimates are significantly different from the estimates obtained by OLS.

Although measures of subjective health, mortality outcomes, and health behaviors are reasonably good proxies for health outcomes, they are still far from having the required properties to be representative as a biomarker. By definition, a biomarker is "a characteristic that is objectively measurable and evaluated as an indicator of normal biological processes, pathological processes, or pharmacologic responses to therapeutic intervention." (Biomarkers Definition Working Groups, 2001, p.91). Certain biomarkers are clinically accepted as the "true" objective measure of an individual's underlying health. For example, reduction of elevated arterial blood pressure has been used for decades by clinicians to reflect the reduction in the stress level and in the incidence of stroke and congestive heart failure, whereas serum cholesterol levels are often used as an indicator of the risk of coronary heart disease (see, for example, Wilson and colleagues, 1998).

Econometric evidence on the relationship between schooling and biomarkers is scarce. A few exceptions have been the seminal studies by Mark Berger and J. Paul Leigh (1989), and David Blanchflower and Andrew Oswald (2008, 2009). These papers are three of the very few studies to estimate the impact of schooling on both systolic and diastolic blood pressure, which are generally viewed as valid biomarkers in the medical literature (see Biomarkers Definition Working Groups, 2001). Using data from the Health and Nutrition Examination Survey, Berger and Leigh estimated blood pressure equations and allowing for self-selection into more years of schooling. Using average real per capita income and expenditures on education in the state in which an individual resided from the year of birth to the age of 6 as their instrumental variables, they found schooling to have a small negative effect on blood pressure: an extra year of schooling reduces both systolic and diastolic blood pressure by approximately -0.6 and -0.2 mmHg, respectively. However, as mentioned previously, their IV estimates obtained by Berger and Leigh are likely to approximate average treatment effects among a subset of population who are responsive to the instruments. Using both the 1998-2007 Health Survey for England and the 2001 Eurobarometer data set, Blanchflower and Oswald find schooling to be negatively correlated with blood pressure, although no attempts have been made to treat schooling as exogenous. Like Berg and Leigh, the schooling coefficients obtained in Blanchflower and Oswald, though statistically significant at the 1% level, are very small and of almost no economic importance. To the best of my knowledge, there have not been any attempts to estimate the

LATE of one extra year of schooling using the compulsory school law as instruments on a biomarker in the UK.

III. Data and empirical strategy

A. *Data*

The data set used in this paper is the Health Survey for England (HSE). The HSE is an annual survey and is designed to monitor the nation's health. The unit of survey in the HSE is the household. Information is collected through a combination of face-to-face interviews, a self-completed questionnaire, and a series of medical examination (including taking measurements for height and weight, as well as recording of blood and saliva sample for clinical tests) conducted by a trained nurse. Three continuous blood pressure measurements are available in 99% of the case, so I take the average for both systolic and diastolic. By definition, systolic blood pressure measures the rate of contraction of heart chambers while driving blood out of the chambers, whereas diastolic blood pressure measures the time when the heart fills with blood after contraction. Both are measured in millimeter of mercury (mmHg). Hypertension is categorized by having a systolic blood pressure of 140 mmHg and above and/or a diastolic blood pressure of 90 mmHg and above.

The schooling variable is recorded as the age an individual finished full-time education. In this paper, I pool data from the 1998 to 2007 HSEs, resulting in the final sample of 75,814 individuals who are aged 16 and over. Some summary of descriptive statistics are given in Table A1 in the appendix.

The legislation on changing the minimum school-leaving age from 14 to 15 was first introduced in the 1944 Education Act, with the first increase implemented in 1947. A further increase in the minimum school-leaving age from 15 to 16 subsequently occurred in 1973. Figure 1 plots the fractions of school leavers at age 14 and 15 before and after both legislations in the pool HSEs. Consistent with Harmon and Walker (the UK Family Expenditure Survey) and Oreopoulos (the UK General Household Survey), we can see that a very high fraction of individuals in the HSE left school at age 14 (or less) before 1947. There is, however, a significant drop in the fraction of school leavers at age 14 in 1947: the portion of school leavers at aged 14 fell from 52% in 1946 to 9% in 1950. A sharp – albeit relatively smaller – drop in the portion of 15 years-old school leavers occurred in 1973: Over the course of three years between 1972 and 1975, the fraction of school-leavers at age 15 (or less) fell from 38% to 15%.

B. *Conventional IV approach*

This paper focuses on one particular prediction made on the demand for health by the human capital model: the effect of schooling on subsequent health outcomes. Following the simple productive efficiency and allocative efficiency frameworks outlined in the previous section, an increase in schooling is predicted to increase the stock of health for the individual, holding other things constant. Let S_i be the number of years formal schooling completed for individual *i*, a health equation with blood pressure, BP_i, as the dependent variable of interest can be written as

$$BP_i = X_i^{'}\gamma + \beta S_i + u_i, \qquad (9)$$

where X is a vector of observable attributes, which may include, among others, age and cohort fixed effects. The parameter β can be thought of as the return to investment in human capital on blood pressure. OLS estimation of equation (9) will yield unbiased estimate of β only if S_i is exogenous. Given that high blood pressure has been used to reflect the increase in the risk of stroke and congestive heart failure, β is expected to take some negative values in the blood pressure equation.

Since schooling is unlikely to be exogenous, I follow prior studies and adopt an IV approach to estimate the LATE for schooling. The first-stage schooling equation can take the following form:

$$S_i = Z_i^{'}\alpha + v_i, \qquad (10)$$

where $Z_i^{'}$ is a vector of variables that are correlated to schooling but are uncorrelated with blood pressure beyond their effects on the endogenous regressor (Angrist and Kruger, 2001). The same principle applies to the estimation of discrete probability models where the outcome variable is a binary variable representing whether the individual has hypertension, cardiovascular problem, or experienced angina, heart attack, or stroke. In these cases, the prefer estimator is the IV-probit (Newey, 1987).

Following Harmon and Walker (1995) and Oreopoulos (2006, 2007), the identification is achieved by the inclusion of dummy variables that record the exogenous change in the minimum school-leaving age law in the UK. In particular, dummy variables are defined for individuals who were aged 14 between 1947 and 1972, and for those who were aged 15 after 1973. The minimum school-leaving age of 14 is the omitted category. Both X and Z include age, age squared, and age cubed to capture the impact of the rate of depreciation in health, gender to capture the gender effect, birth cohort fixed effects to capture the trends in education attainment and health, and regional dummies to capture the geographic-specific effects. All regressions are clustered by birth cohort and regions.

IV. Results

Following Oreopoulos (2006), Figures 2 and 3 provide the graphic illustrations of the compelling effects of minimum school-leaving age on educational attainment and blood pressure using the RD method (see Imbens and Lemieux, 2008). The RD method differs from previous studies by comparing education attainment and subsequent blood pressure just before and after the policy change. Aggregating the data into cell means by birth year and age in order to create cohort averages, Figure 2 plots the average age cohorts left full-time education by the year they were age 14 during 1920 and 1970 (N = 834). The figure also plots the fitted values from regressing the means on age, age-squared, age-cubed, birth cohort dummies, and a dummy for whether or not a cohort faced a minimum school-leaving age of 15. The vertical line represents the change of minimum school-leaving age from 14 to 15 in 1947. Consistent with Harmon and Walker (1995) and Oreopoulos (2006), we can see a clear jump in the average age left full-time education after 1947, with the fit predicts an increase in the schooling level between 1946 and 1947 of 0.4 years (R^2 = 0.899).

Figure 3 plots the corresponding mean systolic blood pressure cohorts using the same sample. Although there is no obvious jump in the mean blood pressure cohorts between 1946 and 1947, we can reject the null hypothesis at the 1% level that the predicted average systolic blood pressure in the years before 1947 (= 148.38 mmHg) and in the years after 1947 (= 136.51 mmHg) is the same. In addition to this, the ratio of the reduced form gradients for these two groups is 0.264 ($-0.168/-0.636$). Figure 3 thus gives the first graphical evidence that, all else equal, blood pressure drops relatively faster in the years that followed the change of minimum school-leaving age from 14 to 15.

Table 1 leaves the analysis of the cohort averages and begins the analysis of the individual-level data by estimating the first-stage OLS schooling regression equation with dummies for compulsory schooling included as the independent variables. Both minimum school-leaving age variables enter the schooling equation positively and statistically significantly at the 1% level. Independent of age-specific, gender, cohort, and regional fixed effects, children who faced the minimum school-leaving age of 15 have approximately 0.4 year more full-time education than those who faced the minimum-schooling age of 14, whereas children who faced the minimum school-leaving age of 16 stayed approximately 0.6

year longer than those who faced the minimum school-leaving age of 14. Note that the 0.4 years jump in the schooling level after 1947 corresponds to the prediction made in Figure 2. The F-statistic [p-value] of the excluded instruments is 36.72 [0.000] and the partial R-squared is 0.0013, which compare favorably with the results reported in Bound, Jaeger, and Baker (1995) of what constitute a weak instrument.

The OLS and IV estimates of the schooling effects on both systolic and diastolic blood pressure are reported in the first four columns of Table 2. Looking at Column 1, we can see from the OLS estimates that there is a negative and statistically significant relationship between schooling and systolic blood pressure, which is consistent with Berger and Leigh (1989), and Blanchflower and Oswald (2008, 2009). However, the correlation is remarkably small, one of almost no economic significance: one additional year of schooling is associated with a decrease in systolic blood pressure of 0.3 mmHg, which is also consistent with the findings in previous studies. The correlation between schooling and diastolic blood pressure is negative though statistically insignificantly different from zero, and an additional year of schooling is associated with a decrease in diastolic blood pressure of only 0.01 mmHg.

By contrast, we can clearly see from Columns 2 and 4 of Table 2 that the IV estimated schooling effects are substantially larger than the ones estimated by OLS. Treating schooling as exogenous, an additional year of schooling reduces systolic blood pressure by approximately 7 mmHg, which is equivalent to an increase of around 2,200% from the OLS estimated schooling effect. Similarly, an additional year of schooling reduces diastolic blood pressure by almost 2 mmHg in the IV equation (or a 200% increase from the estimates obtained by OLS). The t-statistics of both coefficients are large, and we can reject the null hypothesis of zero schooling effects in both blood pressure equations at the 1% level. Note that a similar set of schooling coefficients are obtained if we were to follow simple health equations and replace birth cohorts fixed effects by survey year dummies (see, e.g., Powdthavee, 2009).

The estimates of the schooling effects on blood pressure may be somewhat difficult to grasp. An obvious question is what makes a healthy range of blood pressure readings? A more practical way is then to estimate the effects of schooling on the probability of developing hypertension in adulthood. This is carried out in the last four columns of Table 2 where probit and IV-probit models of systolic and diastolic hypertension equations are estimated. For systolic hypertension, the dependent variable takes the value of 1 if systolic blood pressure is recorded at 140 mmHg or over and 0 otherwise, whereas the dependent variable in the diastolic hypertension takes the value of 1 if diastolic blood pressure is recorded at 90 mmHg or over and 0 otherwise. These ranges are known by clinicians as the ranges representing Stage I hypertension, and individuals with blood pressure within these ranges are normally required medical consultation or a lifestyle change.

Schooling is negatively correlated with the probability of having both systolic and diastolic hypertension in the probit equations, although the effect is statistically significant only in the systolic hypertension equation. By contrast, IV-probit yields the estimated schooling effects that are negative and significant at the 1% level in both sets of hypertension equations. The estimated schooling effects are also quantitatively important as well as statistically significant. For systolic hypertension, the probit coefficients imply that completing an additional year of schooling reduces the probability of having hypertension from 29% to 28% in the probit equation, whereas the estimated IV-probit coefficients imply that completing an additional year of schooling reduces the same probability from 29% to 18%, all else equal. With respect to diastolic hypertension, an additional year of schooling in the IV-probit reduces the probability of an individual having hypertension by 5% points from 9% to 4%. By contrast, the estimated marginal effect of schooling on diastolic hypertension is measured at less than 0.1% point in the probit equation.

Table 3 re-estimates previous table's regression equations on an age group that perhaps face the highest risk of developing cardiovascular diseases, i.e. individuals aged between 35 and 80. This restricts the number of observations by approximately one-quarter of the full sample. The estimated IV schooling coefficients continue to be negative, statistically significant, and markedly larger than those estimated by their OLS (or probit) counterparts. In addition to this, the estimated schooling effects are also slightly larger among the risk group than when the equations are estimated using the full sample.

I introduce in Table 4 some additional controls on socio-economic and demographic status that are not shown explicitly, including one dummy for being married, eight social class dummies, a history of having high blood pressure, six income categories, as well as number of children, number of years smoked, BMI and BMI-squared. The equations are then re-estimated using the full sample and the risk group. Inconsistent with the allocative efficiency model, the inclusion of such additional control variables as BMI and number of years smoked do not lead to a reduction in the size of the schooling effects. For example, the IV coefficients in the full sample estimate imply that completing an addition year of schooling lowers systolic and diastolic blood pressure by approximately 8 mmHg and 3 mmHg, respectively, which are similar to the estimates obtained in Table 2.

Finally, Table 5 reports IV-probit estimates of the schooling effects on the probability of having (i) a cardiovascular condition, and (ii) had incidence of angina, heart-attack, or stroke among those aged 16 and over, as well as the risk group. The binary cardiovascular condition variable takes the value of 1 if the trainee nurse reported that individual has one of the following conditions: irregular heart rhythm; high blood pressure; angina; or heart attack or stroke. The heart attack/stroke, on the other hand, takes the value of 1 if the individual has experienced either a heart attack or a stroke. Looking across the table, we can see that the IV schooling coefficient enters both cardiovascular condition and heart attack/stroke equations in a negative manner. However, the LATE for schooling is only statistically significant in the heart attack/stroke equations. The results are robust to controlling for age, age-squared, age-cubed, birth cohort fixed effects, as well as other socio-economic demographic variables used to estimate hypertension equations in Table 4. Among those aged 16 and over, the IV-probit estimates imply that an additional year of schooling reduces the probability of having a cardiovascular condition from 25% to 21% and angina, heart attack or stroke from 6% to 5%. Similar marginal effects for schooling are obtained for the risk group.

V. Conclusion

This paper is the first of its kind to estimate the biomarker effect of compulsory schooling. Using changes in the minimum school-leaving age law in the UK as instruments in a variety of blood pressure equations, the IV-probit estimates imply that completing an additional year of schooling reduces the probability of developing subsequent hypertension by approximately 5%-11% points. The results are robust to a large set of socio-economic controls, BMI, and years spent smoking, which is consistent with Grossman’s (2005) productive efficiency theory on the direct impact of human capital on the individual’s demand for health. In both IV and IV-probit regressions, the correct estimates of the LATE for schooling indicate the presence of a large and negative bias in the least square/probit estimates of schooling-health relationship. Furthermore, given the compelling effects of the minimum school-leaving age law in the UK – in that virtually every student who was age 14 between 1947 and 1972, and age 15 from 1973 and thereafter were affected by the law, it is likely that my IV estimates of the schooling effect on blood pressure are close to mirroring the ATE for the general population in England (Oreopoulos, 2006).

This paper has important policy implications. First, it provides evidence in favor of Grossman (1975) who has suggested that an increase in expenditure on education rather than on health itself is perhaps the most cost effective way to improve the nation's health. Secondly, according to the Office for National Statistics in the UK, death by heart disease explains around 20% of total death in 2005, which is considerably higher than the second-place killer: cerebrovascular diseases (8%), and the third-place killer: lung cancer (7%). If an additional year of schooling can help reduce the incidence of hypertension by approximately 11% points, then the implications of a nationwide change in the minimum school-leaving age from 16 to 18, which is scheduled to take place in 2013, on the nation's well-being may have been underestimated if one was to simply look at the market returns to education.

The results of this paper also call for further inquiry into the estimation of compulsory schooling on other biomarkers, including, among others, cholesterol and cortisol levels, in order for academics and policy makers to obtain a more complete picture of the relationship between schooling and the true health of a nation.

Reference

Adams, Scott J. 2002. Educational attainment and health: Evidence from a sample of older adults. *Education Economics*, 10, 97-109.

Angrist, Joshua D. 2004. Treatment effect heterogeneity in theory and practice, *Economic Journal*, 114(494), C52-C83.

Angrist, Joshua D., and Kruger, Alan B. 2001. Instrumental variables and the search for identification: From supply to demand and natural experiments. *Journal of Economic Perspectives*, 15(4), 69-85.

Angrist, Joshua D., and Kruger, Alan B. 1991. Does compulsory school attendance affect schooling and earnings? *Quarterly Journal of Economics*, 106(4), 979-1014.

Arendt, Jacob N. 2005. Does education cause better health? A panel data analysis using school reforms for identification. *Economics of Education Review*, 24(2), 149-160.

Behrman, Jere R., and Rosenzweig, Mark R.. 2002. Does increasing women's schooling raise the schooling for the next generation? *American Economic Review*, 92, 323-334.

Berger, Mark C., and Leigh, J. Paul. 1989. Schooling, self-selection, and health. *Journal of Human Resources*, 24(3), 433-455.

Biomarkers Definition Working Groups. 2001. Biomarkers and surrogate endpoints: Preferred definitions and conceptual framework. *Clinical, Pharmacology, & Therapeutics*, 69(3), 89-95.

Blackburn, McKinley L., and Neumark, David. 1995. Are OLS estimates of the return to schooling biased downward? Another look. *Review of Economics and Statistics*, 77(2), 217-229.

Blanchflower, David G., and Oswald, A.J. 2008. Hypertension and happiness across nations. *Journal of Health Economics*, 27(2), 218-233.

Blanchflower, David G., and Oswald, A.J. 2009. Hypertension and happiness across regions. Department of Economics, University of Warwick, mimeo.

Bound, John; Jaeger, David, and Baker, Regina. 1995. Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variable is weak. *Journal of the American Statistical Association*, 90(430), 443-450.

Card, David. 1995. Using geographic variation in college proximity to estimate the return to schooling. In Louis N. Christofides, E. Kenneth Grant, and Robert Swindisky (eds.), *Aspects of Labour Market Behaviour: Essays in Honour of John Vanderkamp*. Toronto: University of Toronto Press, p.201-222.

Card, David. 2001. Estimating the return to schooling: Progress on some persistent econometric problems. *Econometrica*, 69, 1127-1160.

Case, Anne; Fertig, Angela, and Paxson, Christina. 2005. The lasting impact of childhood health and circumstance. *Journal of Health Economics*, 24, 365-389.

Chou, Shin-Yi; Grossman, Michael, and Saffer, Henry. 2004. An economic analysis of adult obesity: Results from the behavioral risk factor surveillance system. *Journal of Health Economics*, 23(3), 565-587.

De Walque, Damien. 2007. How does the impact of an HIV/AIDS information campaign vary with education attainment? *Journal of Development Economics*, 84(2), 686-714.

Erbsland, Manfred; Reid, Walter, and Ulrich, Volker. 1995. Health, health care, and the environment. Econometric evidence from the German micro data. *Health Economics*, 4(3), 169-182.

Fraser, Gary E. 1986. *Preventive Cardiology*. New York: Oxford University Press.

Fuch, Victor R. 1982. Time preference and health: An exploratory study. In Victor R. Fuch (ed), *Economic Aspects of Health*, University of Chicago Press: Chicago, p.93-120.

Gerdtham, Ulf-G., and Johannesson, Magnus. 1999. New estimates of the demand for health: Results based on a categorical health measure and Swedish micro data. *Social Science & Medicine*, 49, 1325-1332.

Goldman, Dana, and Smith, James P. 2002. Can patient self-management help explain the SES health gradient? *Proceedings of the National Academy of Sciences of the United States*, 99, 10929-10934.

Grossman, Michael. 1972. The demand for health: A theoretical and empirical investigation. Columbia University Press for the National Bureau of Economic Research: New York.

Grossman, Michael. 1975. The correlation between health and schooling. In Nestor E. Terleckyj (ed.), *Household Production and Consumption*, Columbia University Press for the National Bureau of Economic Research: New York, p.147-211.

Grossman, Michael. 2000. The human capital model. In Antony J. Culyer and Joseph P. Newhouse (eds.), Handbook of Health Economics, Vol. 1A, Elsevier: Amsterdam, p.347-408.

Grossman, Michael. 2005. Education and nonmarket outcomes. NBER Working Paper 11582.

Harmon, Colm, and Walker, Ian. 1995. Estimates of the return to schooling for the United Kingdom. *American Economic Review*, 85(5), 1278-1286.

Hartog, Joop, and Oosterbeek, Hessel. 1998. Health, wealth, and happiness: Why pursue a higher education? *Economics of Education Review*, 17, 245-256.

Hoffman, Albert; Feinleib, Manning; Garrison, Robert J., and van Laar, Abraham. 1983. Does change in blood pressure predict heart disease? *British Medical Journal*, 287, 267-269.

Imbens, Guido W., and Angrist, Joshua D. 1994. Identification and estimation of local average treatment effects. *Econometrica*, 62(2), 467-475.

Imbens, Guido W., and Lemieux, Thomas. 2008. Regression discontinuity designs: A guide to practice. *Journal of Econometrics*, 142(2), 615-635.

Kenkel, Donald S. 1991. Health behavior, health knowledge, and schooling. *Journal of Political Economy*, 99(2), 287-305.

Leigh, J. Paul, and Dhir, Rachna. 1997. Schooling and frailty among seniors. *Economics of Education Review*, 16(1), 45-57.

Lleras-Muney, Adriana. 2005. The relationship between education and adult mortality in the United States. *Review of Economics Studies*, 72, 189-221.

Michael, Robert T. 1972. *The Effect of Education on Efficiency in Consumption*. Columbia University Press for the National Bureau of Economic Research, New York.

Michael, Robert T. 1973. Education and nonmarket production. *Journal of Political Economy*, 81, 306-327.

Michael, Robert T., and Becker, Gary S. 1973. On the new theory of consumer behavior. *Swedish Economic Journal*, 75, 378-396.

Newey, Whitney K. 1987. Efficient estimation of limited dependent variable models with endogenous explanatory variables. *Journal of Econometrics*, 36(3), 231-250.

Oreopoulos, Philip. 2006. Estimating local and average treatment effects when compulsory schooling really matter. *American Economic Review*, 96(1), 152-175.

Oreopoulos, Philip. 2007. Do dropouts drop out too soon? Wealth, health and happiness from compulsory schooling. *Journal of Public Economics*, 91(11-12), 2213-2229.

Powdthavee, Nattavudh. 2009. Ill-health as a household norm: Evidence from other people's health problems. *Social Science & Medicine*, 68, 251-259.

Rosenzweig, Mark R. 1995. Why are there returns to schooling? *American Economic Review*, 85, 153-158.

Sander, William. 1995. Smoking and quitting smoking. *Review of Economics and Statistics*, 77, 191-199.

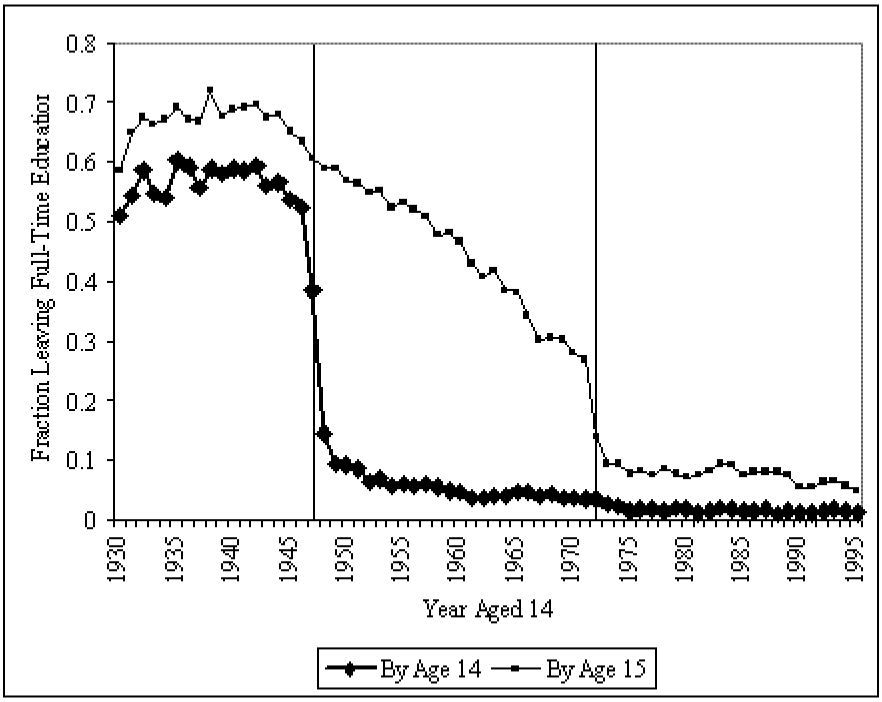
Spasojevic, Jelena. 2003. Effects of education on adult health in Sweden: Results from a natural experiment. PhD Dissertation, City University of New York Graduate Center, New York.

Wagstaff, Adam. 1986. The demand for health: Some new empirical evidence. *Journal of Health Economics*, 5(3), 195-233.

Wilson, Peter W.F., and colleagues. 1998. Prediction of coronary heart disease using risk factor categories. *Circulation*, 97, 1837-1847.

Figure 1: Fraction of Students Left Full-Time Education By Age 14 and 15,

HSE 1998-2007



Note: the lower line represents the proportion of adults (aged 16 and over) in the HSE who left school at or before the age of 14 between 1930 and 1995. The upper line shows the same, but for age 15. The vertical lines indicate the introduction of minimum school-leaving age law in the United Kingdom, which first occurred in 1947 (from age 14 to 15) and then again in 1973 (from age 15 to 16).

Figure 2: Local Averages and Parametric Fit of Average Schooling Age By Year Aged 14

Note: Local averages are plotted for English individuals who were aged 14 between 1920 and 1970. The curved line represents the predicted fit by regressing mean age finished full-time education on age, age-squared, age-cubed, and birth cohort fixed effects. The minimum school-leaving age was raised from age 14 to 15 in 1947.

Figure 3: Local Averages and Parametric Fit of Average Systolic Blood Pressure By Year Aged 14

Note: Local averages are plotted for English individuals who were aged 14 between 1920 and 1970. The curved line represents the predicted fit by regressing mean systolic blood pressure on age, age-squared, age-cubed, and birth cohort fixed effects. The minimum school-leaving age was raised from age 14 to 15 in 1947.

Table 1: Least Squares First-Stage Schooling Equation,

Health Survey for England 1998-2007

Dependent variable: Age finished full-time education	Coefficient
Minimum school-leaving age = 15	0.364
	[0.054]**
Minimum school-leaving age = 16	0.571
	[0.067]**
Age	0.186
	[0.029]**
Age-squared	-0.003
	[0.000]**
Age-cubed	0.00001
	[0.000]**
Female	-0.057
	[0.013]**
Born: 1920-1924	0.074
	[0.067]
Born: 1925-1929	0.217
	[0.079]**
Born: 1930-1934	0.320
	[0.087]**
Born: 1935-1939	0.617
	[0.112]**
Born: 1940-1944	0.815
	[0.115]**
Born: 1945-1949	1.061
	[0.122]**
Born: 1950-1954	1.378
	[0.128]**
Born: 1955-1959	1.524
	[0.132]**
Born: 1960-1964	1.740
	[0.135]**
Born: 1965-1969	2.038
	[0.141]**
Born: 1970-1974	2.407
	[0.145]**
Born: 1975-1979	2.999
	[0.158]**
Born: 1980-1984	2.943
	[0.176]**

Born: 1985-1989	2.683
	[0.207]**
Born: 1990+	2.165
	[0.228]**
North West and Merseyside	0.200
	[0.035]**
Yorkshire and the Humberside	0.169
	[0.038]**
West Midlands	0.198
	[0.033]**
East Midlands	0.219
	[0.035]**
Eastern	0.367
	[0.035]**
London	0.864
	[0.040]**
South East	0.624
	[0.034]**
South West	0.478
	[0.036]**
Constant	10.616
	[0.519]**
Observations	75,814
R-squared	0.2514
Partial R-squared of excluded instruments	0.0013
Test of excluded instruments: F(2, 147)	36.72
	[0.000]**

Note: +<10%; *<5%; **<1%. Standard errors are in parentheses. Standard errors are clustered by birth cohort and region.

Table 2: OLS, IV, Probit, and IV-Probit Blood Pressure and Hypertension Equations, Health Survey for England 1998-2007

	Systolic		Diastolic		Hypertension (Systolic)		Hypertension (Diastolic)	
	OLS	IV	OLS	IV	Probit	IV-Probit	Probit	IV-Probit
Age finished full-time education	-0.331	-7.383	-0.011	-1.977	-0.028	-0.343	-0.006	-0.279
	[0.044]**	[1.819]**	[0.029]	[0.791]*	[0.003]**	[0.066]**	[0.004]	[0.115]*
Age	-1.079	0.299	2.273	2.658	-0.041	0.026	0.360	0.384
	[0.339]**	[0.471]	[0.247]**	[0.271]**	[0.023]+	[0.024]	[0.036]**	[0.029]**
Age-squared	0.015	-0.007	-0.031	-0.037	0.001	0.0001	-0.006	-0.006
	[0.008]+	[0.009]	[0.006]**	[0.006]**	[0.000]	[0.000]	[0.001]**	[0.001]**
Age-cubed	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
	[0.000]+	[0.000]	[0.000]**	[0.000]**	[0.000]	[0.000]	[0.000]**	[0.000]**
Female	-5.045	-5.448	-2.851	-2.964	-0.264	-0.252	-0.266	-0.057
	[0.488]**	[0.514]**	[0.172]**	[0.181]**	[0.029]**	[0.026]**	[0.018]**	[0.013]**
Born: 1920-1924	-7.064	-6.645	-4.621	-4.502	-0.261	-0.212	-0.374	-0.326
	[1.311]**	[1.124]**	[0.984]**	[0.924]**	[0.071]**	[0.058]**	[0.078]**	[0.072]**
Born: 1925-1929	-11.382	-10.046	-6.674	-6.300	-0.477	-0.362	-0.616	-0.512
	[1.772]**	[1.527]**	[1.321]**	[1.222]**	[0.093]**	[0.080]**	[0.101]**	[0.078]**
Born: 1930-1934	-17.096	-14.090	-9.179	-8.336	-0.738	-0.517	-0.853	-0.664
	[2.181]**	[2.030]**	[1.572]**	[1.450]**	[0.113]**	[0.115]**	[0.117]**	[0.138]**
Born: 1935-1939	-21.796	-15.305	-10.839	-9.025	-0.965	-0.561	-1.080	0.614
	[2.412]**	[2.687]**	[1.735]**	[1.688]**	[0.123]**	[0.148]**	[0.130]**	[0.109]**
Born: 1940-1944	-27.151	-19.404	-12.121	-9.958	-1.271	-0.774	-1.276	0.812
	[2.548]**	[3.045]**	[1.820]**	[1.806]**	[0.130]**	[0.176]**	[0.136]**	[0.113]**
Born: 1945-1949	-31.924	-22.580	-13.238	-10.631	-1.517	-0.920	-1.438	-0.955

	[2.567]**	[3.437]**	[1.831]**	[1.889]**	[0.132]**	[0.203]**	[0.138]**	[0.120]**
Born: 1950-1954	-37.155	-25.701	-14.547	-11.351	-1.821	-1.093	-1.542	1.374
	[2.532]**	[3.917]**	[1.801]**	[1.985]**	[0.134]**	[0.237]**	[0.140]**	[0.315]**
Born: 1955-1959	-42.378	-29.384	-15.716	-12.092	-2.155	-1.320	-1.660	-1.018
	[2.482]**	[4.316]**	[1.776]**	[2.057]**	[0.134]**	[0.267]**	[0.145]**	[0.348]**
Born: 1960-1964	-46.902	-31.647	-16.669	-12.415	-2.431	-1.461	-1.754	1.744
	[2.425]**	[4.829]**	[1.734]**	[2.199]**	[0.134]**	[0.303]**	[0.145]**	[0.388]**
Born: 1965-1969	-50.428	-33.136	-16.848	-12.027	-2.667	-1.578	-1.778	2.042
	[2.413]**	[5.327]**	[1.724]**	[2.353]**	[0.137]**	[0.334]**	[0.149]**	[0.422]*
Born: 1970-1974	-52.884	-33.029	-15.516	-9.980	-2.778	-1.560	-1.580	-0.677
	[2.434]**	[5.956]**	[1.744]**	[2.579]**	[0.144]**	[0.366]**	[0.165]**	[0.145]**
Born: 1975-1979	-55.813	-31.788	-14.103	-7.404	-2.932	-1.508	-1.460	3.002
	[2.531]**	[6.992]**	[1.809]**	[2.976]*	[0.155]**	[0.413]**	[0.176]**	[0.509]
Born: 1980-1984	-57.740	-34.081	-12.006	-5.409	-3.015	-1.598	-0.992	0.010
	[2.665]**	[6.931]**	[1.924]**	[2.994]+	[0.172]**	[0.418]**	[0.215]**	[0.176]**
Born: 1985-1989	-62.558	-40.703	-9.736	-3.641	-3.401	-2.021	-0.802	0.114
	[2.869]**	[6.615]**	[2.060]**	[2.927]	[0.189]**	[0.427]**	[0.258]**	[0.463]
Born: 1990+	-67.374	-49.137	-8.937	-3.850	-3.513	-2.284		
	[3.463]**	[6.011]**	[2.572]**	[2.961]	[0.398]**	[0.503]**		
North West and Merseyside	-1.693	-0.280	-0.935	-0.541	-0.086	-0.012	-0.049	0.200
	[0.252]**	[0.453]	[0.241]**	[0.263]*	[0.025]**	[0.026]	[0.033]	[0.035]**
Yorkshire and the Humberside	-0.367	0.828	-0.010	0.323	0.025	0.076	-0.084	-0.031
	[0.250]	[0.466]+	[0.257]	[0.277]	[0.025]	[0.023]**	[0.033]*	[0.039]
West Midlands	-1.619	-0.218	-1.319	-0.929	-0.089	-0.015	-0.066	0.198
	[0.271]**	[0.475]	[0.241]**	[0.276]**	[0.022]**	[0.024]	[0.032]*	[0.033]**
East Midlands	-1.972	-0.420	-1.464	-1.031	-0.119	-0.036	-0.062	0.219

	[0.314]**	[0.507]	[0.257]**	[0.293]**	[0.028]**	[0.028]	[0.039]	[0.035]**
Eastern	-2.828	-0.245	-1.884	-1.164	-0.157	-0.022	-0.157	-0.043
	[0.292]**	[0.768]	[0.226]**	[0.347]**	[0.025]**	[0.037]	[0.030]**	[0.059]
London	-3.938	2.156	-1.791	-0.093	-0.205	0.094	-0.108	0.864
	[0.320]**	[1.655]	[0.225]**	[0.701]	[0.029]**	[0.075]	[0.035]**	[0.113]
South East	-2.226	2.171	-1.667	-0.441	-0.127	0.086	-0.108	0.624
	[0.246]**	[1.163]+	[0.249]**	[0.529]	[0.024]**	[0.051]+	[0.031]**	[0.034]**
South West	-2.187	1.200	-1.402	-0.458	-0.114	0.052	-0.059	0.078
	[0.264]**	[0.930]	[0.241]**	[0.435]	[0.027]**	[0.045]	[0.035]+	[0.068]
Constant	213.861	289.124	49.644	70.610	3.338	6.354	-6.082	-2.666
	[6.209]**	[20.659]**	[4.469]**	[9.894]**	[0.433]**	[0.672]**	[0.672]**	[1.806]
Observations	75,814	75,814	75,811	75,811	75,814	75,814	75,776	75,776
R-squared	0.240		0.100		0.157		0.065	
Hanson J Statistic		1.660		0.022				
(Over-identification)		[0.1977]		[0.8811]				

Note: See Table 1. Systolic hypertension = 1 if systolic blood pressure ≥ 140 mmHg, and 0 otherwise. Diastolic hypertension = 1 if diastolic blood pressure ≥ 90 mmHg, and 0 otherwise

Table 3: OLS, IV, Probit, and IV-Probit Blood Pressure and Hypertension Equations: Risk Group (35<=Age<=80)

	Systolic		Diastolic		Hypertension (Systolic)		Hypertension (Diastolic)	
	OLS	IV	OLS	IV	Probit	IV-Probit	Probit	IV-Probit
Age finished full-time education	-0.409	-7.646	-0.055	-2.448	-0.030	-0.351	-0.006	-0.327
	[0.052]**	[1.691]**	[0.035]	[0.850]**	[0.004]**	[0.057]**	[0.005]	[0.112]**
Observations	54200	54200	54198	54198	54200	54200	54198	54198
R-squared	0.19		0.05		0.119		0.0339	
Hanson J Statistic		0.009		3.929				
(Over-identification)		[0.9259]		[0.0475]				

Note: +<10%; *<5%; **<1%. Standard errors are in parentheses. Standard errors are clustered by birth cohort and region. Same control variables as in Table 1.

Table 4: IV and IV-Probit Blood Pressure Equations with Socio-Economic and Demographic Controls

1) Full sample (age 16+)	IV	IV	IV-Probit	IV-Probit
Dependent variable	Systolic Blood Pressure	Diastolic Blood Pressure	Hypertension (Systolic)	Hypertension (Diastolic)
Age finished full-time education	-8.444	-2.546	-0.451	-0.357
	[2.298]**	[1.271]*	[0.079]**	[0.140]*
2) Risk group (35<=age<=80)	IV	IV	IV-Probit	IV-Probit
Dependent variable	Systolic Blood Pressure	Diastolic Blood Pressure	Hypertension (Systolic)	Hypertension (Diastolic)
Age finished full-time education	-9.039	-3.319	-0.400	-0.455
	[2.240]**	[1.255]**	[0.125]**	[0.071]**

Note: +<10%; *<5%; **<1%. Standard errors are in parentheses. Standard errors are clustered by birth cohort and region. Controls include age, age-squared, birth cohort fixed effects, eight regional dummies, one dummy for being married, eight social class dummies, a history of having high blood pressure, six income categories, as well as number of children, number of years smoked, BMI and BMI-squared.

Table 5: IV-Probit Cardiovascular Problems Equations with Socio-Economic and Demographic Controls

1) Full sample (age 16+)	IV-Probit	IV-Probit
Dependent variable	Cardiovascular conditions	Angina/heart attack/stroke
Age finished full-time education	-0.235	-0.217
	[0.145]	[0.108]*
Observation	51903	63602
2) Risk group (35<=age<=80)	IV-Probit	IV-Probit
Dependent variable	Cardiovascular conditions	Angina/heart attack/stroke
Age finished full-time education	-0.169	-0.195
	[0.152]	[0.104]+
Observation	35453	45432

Note: Same controls as in Table 3.

Table 1A: Summary Statistics

	All		Left school age 14		Left school age 15		Left school age 16	
	M	SD	M	SD	M	SD	M	SD
Systolic blood pressure	132.676	19.567	147.662	24.350	136.021	18.491	124.282	13.557
Diastolic blood pressure	74.548	12.463	75.249	16.744	77.452	11.620	71.669	10.623
Hypertension (systolic > 140mmHG)	0.299	0.458	0.610	0.488	0.372	0.483	0.121	0.326
Hypertension (diastolic > 90mmHG)	0.096	0.294	0.123	0.329	0.134	0.340	0.052	0.221
Age finished full-time education	16.439	1.689	15.011	1.547	16.211	1.580	17.157	1.431
Minimum school-leaving age = 15	0.400	0.490						
Minimum school-leaving age = 16	0.442	0.497						
Age	47.779	19.508	77.557	6.312	56.607	7.762	32.889	7.771
Female	0.547	0.498	0.569	0.495	0.532	0.499	0.553	0.497
Born: 1920-1924	0.032	0.175	0.201	0.400				
Born: 1925-1929	0.051	0.219	0.320	0.466				
Born: 1930-1934	0.062	0.241	0.231	0.421	0.064	0.245		
Born: 1935-1939	0.075	0.263			0.187	0.390		
Born: 1940-1944	0.074	0.261			0.184	0.388		
Born: 1945-1949	0.090	0.286			0.225	0.417		
Born: 1950-1954	0.084	0.278			0.211	0.408		
Born: 1955-1959	0.090	0.286			0.129	0.335	0.087	0.281
Born: 1960-1964	0.105	0.306					0.237	0.425
Born: 1965-1969	0.099	0.299					0.225	0.418
Born: 1970-1974	0.082	0.275					0.186	0.389
Born: 1975-1979	0.061	0.239					0.138	0.345
Born: 1980-1984	0.044	0.206					0.100	0.300

Born: 1985-1989	0.011	0.107					0.026	0.159
Born: 1990+	0.000	0.021					0.001	0.032
North West and Merseyside	0.139	0.346	0.130	0.337	0.140	0.347	0.142	0.349
Yorkshire and the Humberside	0.103	0.303	0.094	0.292	0.104	0.305	0.105	0.306
West Midlands	0.101	0.302	0.094	0.292	0.103	0.304	0.102	0.303
East Midlands	0.098	0.298	0.092	0.290	0.098	0.298	0.100	0.300
Eastern	0.110	0.313	0.101	0.301	0.117	0.321	0.107	0.309
London	0.126	0.332	0.092	0.289	0.113	0.316	0.150	0.357
South East	0.141	0.348	0.145	0.352	0.147	0.354	0.134	0.341
South West	0.111	0.314	0.120	0.325	0.117	0.321	0.103	0.304
Observations	75,814		11,997		30,320		33,497	